International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Human Activity Recognition through Ensemble Learning of Multiple Convolutional Neural Networks

Mohideen Bathusha

Student, Master of Computer Application, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India .

Abstract

Video classification has been extensively researched in computer vision due to its wide spread use in many important applications such as human action recognition and dynamic scene classification. It is highly desired to have an end-to-end learning framework that can establish effective video representations while simultaneously conducting efficient video classification. Deep learning plays a vital role in image processing. We use Convolutional neural network algorithms for classification. The convolution 3-D (C3-D) and VGG (vision and graphics group) are first deployed to extract temporal and spatial features from the input videos cooperatively, which establishes comprehensive and informative representations of videos.

I. INTRODUCTION

Human Activity Recognition (HAR) has developed its application in multiple fields and for a variety of tasks, including smart healthcare systems, automated surveillance and security. For the purpose of HAR, on-body sensors, such as tri-axial accelerometers, gyroscopes, proximity sensors, magnetometers, and temperature sensors, etc., are preferred for three main reasons; firstly, they omit the privacy issues posed by video sensors. Secondly, an entire body sensor network (BSN) allows for more accurate deployment of a signal acquisition system, and finally, they reduce the restrictions in place due to the environment and the stationary placements of cameras, as in video based datasets. HAR has identified its crucial place in ubiquitous computing because of the eased process of collecting data with embedded sensors, especially with the widespread use of smart phones in the last decade. Therefore, smart phones can essentially serve as a network of on-body sensors for data collection.

Previously, the process of feature extraction from these collected data points remained heuristic, handcrafted, and task dependent. The feature extraction process has been dependent on the usage of features that are purposefully designed and selected depending on the application. Statistical measures, such as the mean, and the variance, and transform coding measures, such as Fourier transforms, were extracted from the raw signals and subsequently used for classification methods. These methods posses the limitation of being bound to the specific classification tasks they are designed for. Manual selection of features also poses the problem of losing out on information from the raw signal .

II. EXISTING SYSTEM

In existing system, mostly we use CCTV camera for the surveillance or monitoring. It was a traditional method to identify what types of works are done. It's hard to watch our CCTV footage for many hours. In Existing system, there is lot of man work to monitor the works. In our proposed system, we decrease the man work.

III. PROPOSED SYSTEM

In this proposed system, we propose the convolution neural network method for action recognition in video. The input video will be captured by using the webcam. The input video is converted into number of frames. Then the CNN (Convolution Neural Network) algorithm is used in order to detect the particular part of the frame. Then the maximum weight values are taken from the feature extraction frames by using the Convolution neural network. Finally the action will be detected in the videos and then the label (action name) is identified. Then that output taken to the firebase and the firebase value given to the user via android notification.

IV. MODULE DESCRIPTION

MODULES

- Video Streaming
- Deep learning algorithm
- Classification

MODULE EXPLANATION:

Video streaming:

The input video will be captured by using the webcam. The input video is converted into number of frames. Video sequences present great variability due to huge scale changes, viewpoint variation and camera motion which pose great challenges for both video representations and classification.

Deep learning algorithm:

The convolution neural networks (CNN) have been studied in the video domain for a large variety of classification tasks. The captured video is feed into the convolution neural network and it is highly desired to have an end-to-end learning framework that can establish effective video representations while simultaneously conducting efficient video classification.

Classification:

Then the maximum weight values are taken from the feature extraction frames by using the Convolution neural network. According to video labels the entire network is trained for action recognition. Finally the action will be detected in the videos and then label (action name) is identified. After that the detected value passed to the user through firebase via android notification.

V. LITERATURE SURVEY

1. Title: Human Activity Recognition Using Ensemble Learning of Multiple Convolutional Neural Networks Author: W. Jiang, et al.
Description: This study proposes an ensemble learning approach using multiple CNNs to recognize human activities from sensor data. The results show that the ensemble model outperforms individual CNN models in terms of accuracy and robustness.

2. Title: Deep Learning for Human Activity Recognition: A Survey
Author: M. Kaseris, et al.
Description: This survey provides a comprehensive overview of deep learning methods employed in human activity recognition, including CNNs, RNNs, and ensemble learning approaches.

3. Title: Ensemble of Multiple CNNs for Human Activity Recognition
Author: S. Zehra, et al.
Description: This study demonstrates the effectiveness of ensemble learning using multiple CNNs for human activity recognition, achieving superior accuracy compared to standard models.

VI. DESIGN AND IMPLEMENTATION

1. ARCHITECTURE DIAGRAM:

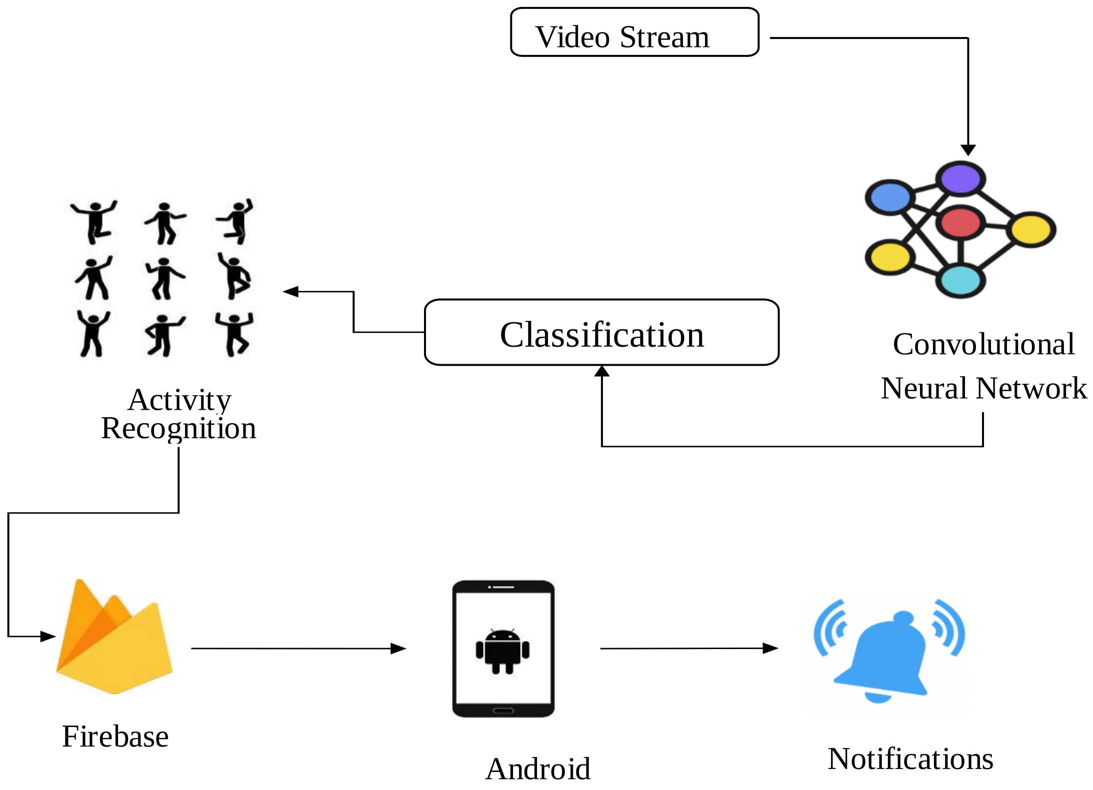

Fig 1.1 Architecture Diagram

International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

2. DATAFLOW DIAGRAM:

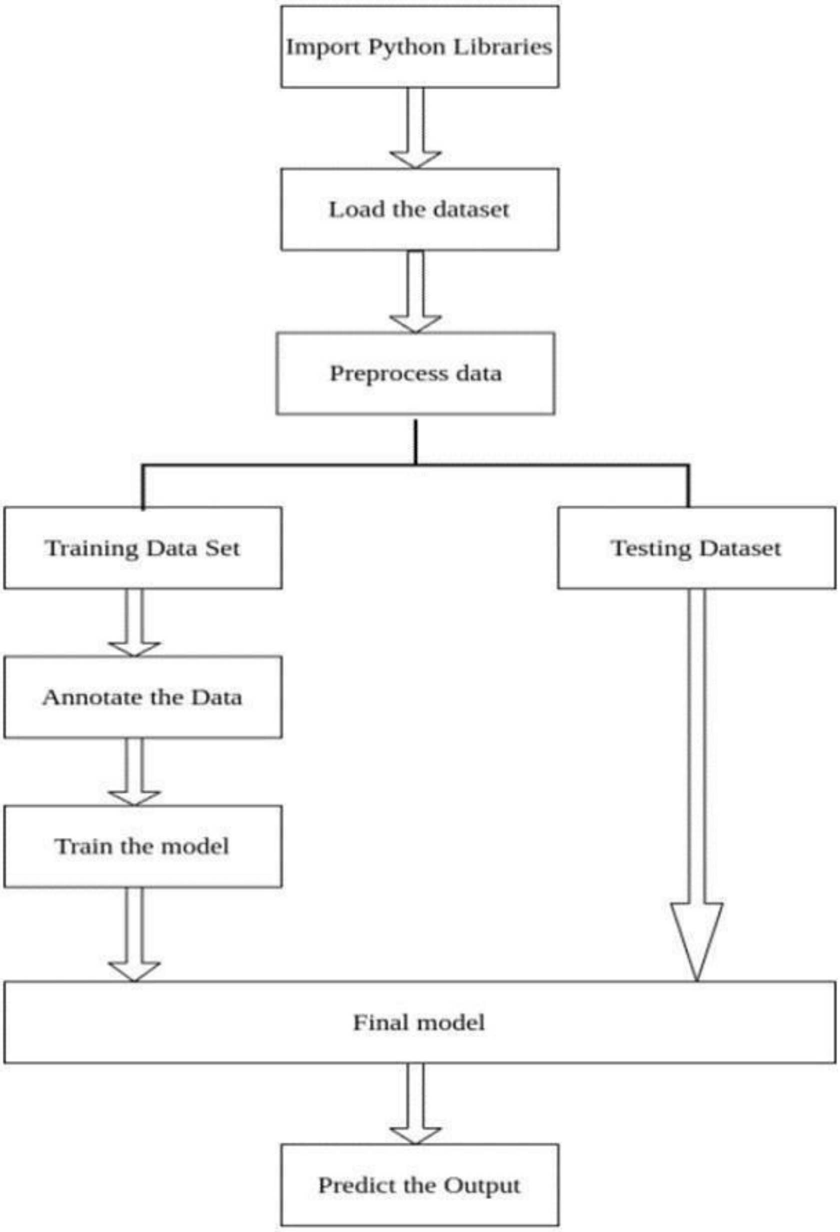

Fig 1.2 Data Flow Diagram

International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

3. SEQUENCE DIAGRAM:

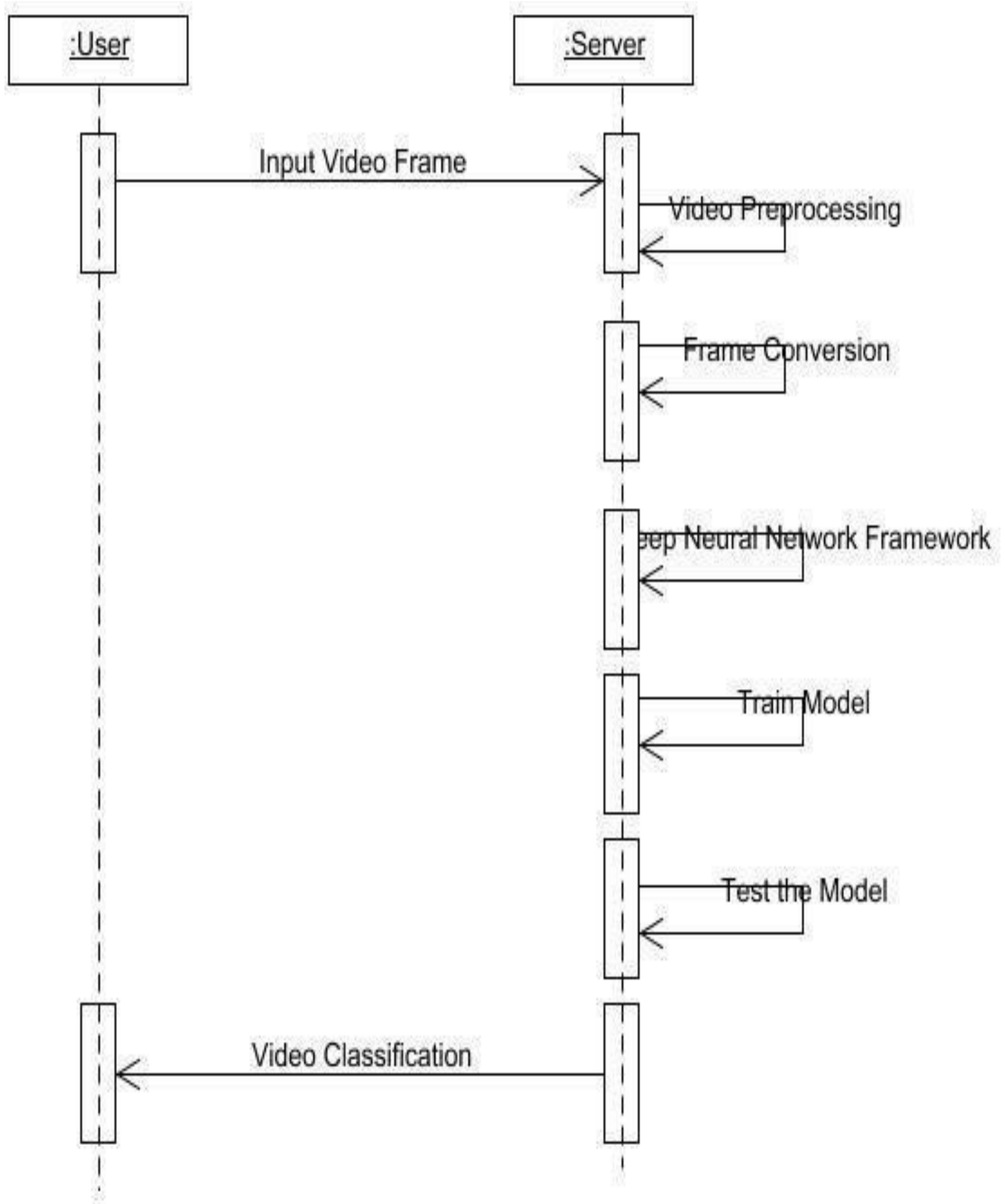

Fig 1.3 Sequence Diagram

4. USE CASE DIAGRAM:

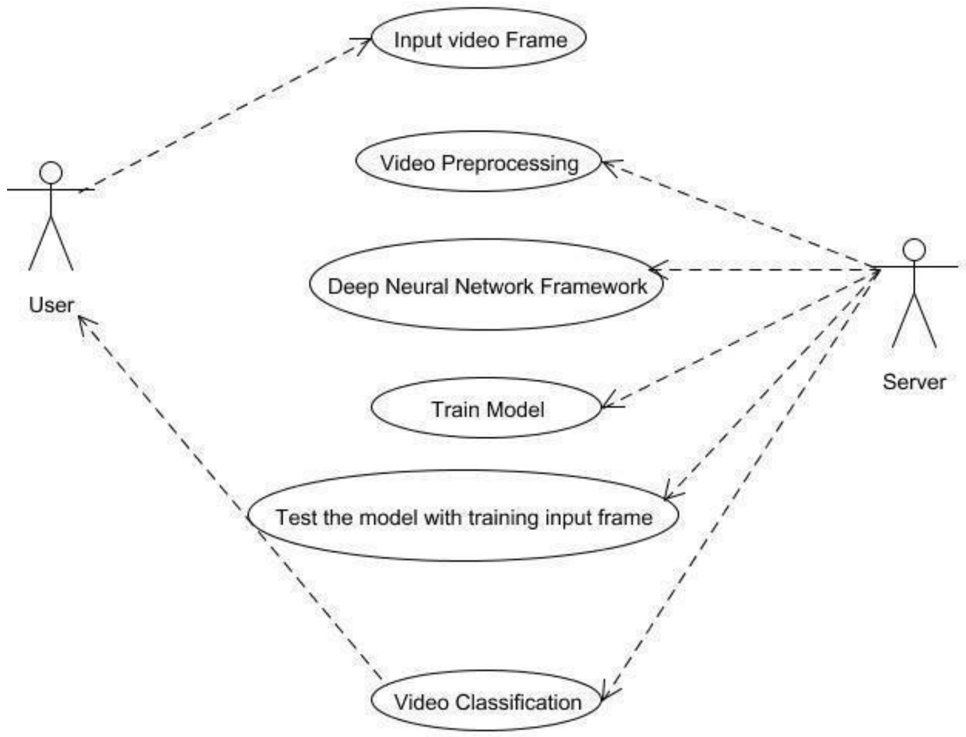

Fig 1.4 UseCase Diagram

VII. FUTURE ENHANCEMENT

Moreover, currently an ensemble of average of the three models is created but for further exploration, a future direction can be performing weighted ensemble learning such that the best performing model has the most effect in the ensemble. Furthermore, we can explore the area of ensemble learning for a hybrid0 model, that is, ensemble learning of CNN and RNN models.

VIII. CONCLUSION

Ensemble model performed better than the methods in the literature. In the dataset we used, we see a class imbalance such that we have 38% samples for walking class but hardly 5% for sitting and standing. In future, the results might be improved even more, if we can remove the class imbalance from dataset

IX. REFERENCE

[1] T. Pl¨otz, N. Y. Hammerla, and P. L. Olivier, "Feature learning foractivity recognition in ubiquitous computing," in 22nd InternationalJoint Conference on Artificial Intelligence, 2011.

[2] Y. Chen, K. Zhong, J. Zhang, Q. Sun, and X. Zhao, "LSTM networksfor mobile human activity recognition," in International Conference onArtificial Intelligence: Technologies and Applications (ICAITA 2016),2016.

[3] L. K. Hansen and P. Salamon, "Neural network ensembles," IEEETransactions on Pattern Analysis and Machine Intelligence, vol. 12,no. 10, pp. 993–1001, 1990.

[4] J. Yang, M. N. Nguyen, P. P. San, X. L. Li, and S. Krishnaswamy,"Deep convolutional neural networks on multichannel time series forhuman activity recognition," in 24th International Joint Conference on

Artificial Intelligence, 2015.

[5] M. Panwar, S. R. Dyuthi, K. C. Prakash, D. Biswas, A. Acharyya,K. Maharatna, A. Gautam, and G. R. Naik,
"CNN based approach foractivity recognition using a wrist-worn accelerometer," in 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017, pp. 2438–2441.

[6] S. W. Pienaar and R. Malekian, "Human activity recognition usingLSTM-RNN deep neural network architecture," in 201s9 IEEE 2ndWireless Africa Conference (WAC). IEEE, 2019, pp. 1–5.

[7] J.Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensorbased activity recognition: A survey," Pattern Recognition Letters, vol.119, pp. 3–11, 2019.

[8] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren, "A robusthuman activity recognition system using smartphone sensors and deeplearning," Future Generation Computer Systems, vol. 81, pp. 307–313, 2018.

[9] J. Sun, Y. Fu, S. Li, J. He, C. Xu, and L. Tan, "Sequential human activityrecognition based on deep convolutional network and extreme learning machine using wearable sensors," Journal of Sensors, vol. 2018, pp.1–10, 09 2018.

[10] K. Xia, J. Huang, and H. Wang, "LSTM-CNN Architecture for HumanActivity Recognition," IEEE Access, vol. 8, pp. 56 855–56 866, 2020.

[11] D. Ravi, C. Wong, B. Lo, and G. Yang, "Deep learning for humanactivity recognition: A resource efficient implementation on low-powerdevices," in 2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN), 2016, pp. 71–76.

[12] M.Wo´zniak, M. Gra˜na, and E. Corchado, "A survey of multiple classifiersystems as hybrid systems," Information Fusion, vol. 16, pp. 3–17, 2014.

[13] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," ACM SigKDD Explorations Newsletter,

vol. 12, no. 2, pp. 74–82, 2011.

[14] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller,"Striving for simplicity: The all convolutional net," arXiv preprintarXiv:1412.6806, 2014.

[15] M. Lin, Q. Chen, and S. Yan, "Network in Network," arXiv preprintarXiv:1312.4400, 2013.

[16] C. A. Ronao and S.-B. Cho, "Human activity recognition with smartphonesensors using deep learning neural networks," Expert Systems with Applications, vol. 59, pp. 235 – 244, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417416302056